2UP

NASA CR- 134223
ERIM 31650-151-T

(NASA-CR-134223) INVESTIGATIONS IN ADAPTIVE PROCESSING OF MULTISPECTRAL DATA Technical Report, 1 Nov. 1971 - 31 Jan. 1973 (Environmental Research Inst. of Michigan) ~~36~~ 37 p HC $5.00 CSCL 05B G3/34

N74-20624

Unclas
35026

INVESTIGATIONS IN ADAPTIVE PROCESSING OF MULTISPECTRAL DATA

by

F. J. Kriegler and H. M. Horwitz
Infrared and Optics Division

ENVIRONMENTAL
RESEARCH INSTITUTE
OF MICHIGAN
FORMERLY WILLOW RUN LABORATORIES.
THE UNIVERSITY OF MICHIGAN

prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Lyndon B. Johnson Space Center
Contract NAS 9-9784, Task B 2.12

NOTICES

Sponsorship. The work reported herein was conducted by the Environmental Research Institute of Michigan for the National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Earth Observations Division, under Contract NAS 9-9784, Task B2.12. Technical Monitor is Dr. Andrew Potter (TF3). Contracts and grants to the Institute for the support of sponsored research are administered through the Office of Contracts Administration.

Disclaimers. This report was prepared as an account of Government-sponsored work. Neither the United States, nor the National Aeronautics and Space Administration (NASA), nor any person acting on behalf of NASA:

(A) Makes any warranty or representation, expressed or implied with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

(B) Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used above, "person acting on behalf of NASA" includes any employee or contractor of NASA, or employee of such contractor, to the extent that such employee or contractor of NASA or employee of such contractor prepares, disseminates, or provides access to any information pursuant to his employment or contract with NASA, or his employment with such contractor.

Availability Notice. Requests for copies of this report should be referred to:

National Aeronautics and Space Administration
Scientific and Technical Information Facility
P. O. Box 33
College Park, Maryland 20740

Final Disposition. After this document has served its purpose, it may be destroyed. Please do not return it to the Environmental Research Institute of Michigan.

NASA CR-
ERIM 31650-151-T

Technical Report

INVESTIGATIONS IN ADAPTIVE PROCESSING OF MULTISPECTRAL DATA

by

F. J. Kriegler and H. M. Horwitz
Infrared and Optics Division

ENVIRONMENTAL
RESEARCH INSTITUTE
OF MICHIGAN
FORMERLY WILLOW RUN LABORATORIES,
THE UNIVERSITY OF MICHIGAN

prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

August 1973

Contract NAS 9-9784

Lyndon B. Johnson Space Center
Houston, Texas 77058

i

FOREWORD

This report describes part of a comprehensive and continuing program of research concerned with advancing the state-of-the-art in remote sensing of the environment from aircraft and satellites. The research is being carried out for the NASA Lyndon B. Johnson Space Center, Houston, Texas, by the Environmental Research Institute of Michigan (ERIM), formerly the Willow Run Laboratories of The University of Michigan. The basic objective of this multidisciplinary program is to develop remote sensing as a practical tool to provide the planner and decision-maker with extensive information quickly and economically.

Timely information obtained by remote sensing can be important to such people as the farmer, the city planner, the conservationist, and others concerned with problems such as crop yield and disease, urban land studies and development, water pollution, and forest management. The scope of our program includes: (1) extending the understanding of basic processes; (2) discovering new applications, developing advanced remote-sensing systems, and improving automatic data processing to extract information in a useful form; and also (3) assisting in data collection, processing, analysis, and ground-truth verification.

The research described here was performed under NASA Contract NAS 9-9784, Task B2.12, and covers the period from November 1, 1971 through January 31, 1973. Dr. Andrew Potter has been Technical Monitor. The program was directed by R. R. Legault, Associate Director of ERIM, and by J. D. Erickson, Principal Investigator and Head of the Multispectral Analysis Section. The ERIM number for this report is 31650-151-T.

ACKNOWLEDGMENTS

The authors wish to acknowledge the direction provided by Mr. R. R. Legault and Dr. J. D. Erickson. Contributions were made by R. E. Marshall, M. Gordon, A. Sato, W. Richardson, R. B. Crane, R. J. Kauth, and P. D. Hyde.

ABSTRACT

Adaptive data processing procedures are applied to the problem of classifying objects in a scene scanned by multispectral sensor. These procedures show a performance improvement over standard nonadaptive techniques. Some sources of error in classification are identified and those correctable by adaptive processing are discussed. Experiments in adaptation of signature means by decision-directed methods are described. Some of these methods assume correlation between the trajectories of different signature means; for others this assumption is not made.

CONTENTS

PRECEDING PAGE BLANK NOT FILMED

FIGURES

INVESTIGATIONS IN ADAPTIVE PROCESSING OF MULTISPECTRAL DATA

1
SUMMARY

Adaptive procedures whereby object signatures may be modified during the classification of multispectral data were investigated. The recognition performances of these procedures were compared with each other as well as with that of a typical nonadaptive classification scheme. The adaptive procedures improved recognition performance significantly.

The adaptive procedures tested modified only the means of the signature distributions, and were decision-directed in that the classification of a data point affected the means modification. In one set of procedures (the noninteractive), only the mean of the recognized data point was modified; in another set (interactive) all of the means were updated on the basis of a single recognized point under the assumption that the trajectories of the means are correlated. Exponentially weighted running estimates of signature means as well as posterior probability weighted estimates of signature means were tested in both noninteractive and interactive versions. The noninteractive procedures showed some superiority; however, this comparative result may have resulted from an oversimplified correlation of the signature trajectories assumed for the interactive procedure.

Sources of error in the classification of multispectral data are discussed; those types best reduced by preprocessing transformations are named as well as those best minimized by adaptive processing. Potentially rewarding paths for continued effort in adaptive multispectral data processing are recommended.

2
INTRODUCTION

Among various remote sensing techniques developed for the purpose of automatically detecting materials on the Earth's surface, one of the more useful is that of multispectral scanning. Typically, an aircraft flying over the area to be sensed carries an imaging scanner. Picking up radiance from the scene, the scanner transmits its output to a spectrometer capable of resolving this radiance into many spectral bands between the ultraviolet and the thermal infrared. After detection, these bands of radiant energy may be recorded as data for subsequent processing by an investigator. By processing the data, we mean applying procedures, algorithms, and computations to the detector outputs in order to provide useful information. For example, an investigator may want to identify the various crop types, their condition, and the total acreage planted to each type. This is often done in digital or analog processing by using a likelihood ratio classification procedure in which fixed parameters for several signature distributions are determined from training sets and extrapolated (or interpolated) from these sets over the complete scene.

Sources of difficulty and error in multispectral data analysis stem mainly from natural variabilities normally found not only in scene composition, but also in the atmosphere and in the incident solar irradiance. For purposes of analysis these sources may be grouped into those encountered (1) along the direction of scan and (2) along the flight path. Along the direction of scan, changes in topography, the sun position, scanner look angle, and atmospheric scattering are principal causes of radiance variations. Along the flight line, more slowly varying functions of time predominate—principally instrument drift, changes in solar elevation, changing meteorological conditions and, over large distances, the changes in latitude (and even season).

2.1. NONADAPTIVE APPROACHES

Toward minimizing the effects of these error sources, we have tried several approaches. The problem of variability along the scan line has been attacked by constructing additive and multiplicative functions fitted to the observed effects; guidance has been found in modeling studies of the atmosphere and in data on the scattering characteristics of typical objects in the scene [1]. This approach has frequently reduced previously experienced variations along the scan line by factors of from 2 to 4, which often allows signature distributions to be separated more effectively. Drift errors have been corrected by controlling gain and offset as functions of scanner-contained calibration sources.

Variability in the scene, however, remains a substantial cause of error, and arises from several sources. In the first place, the objects to be mapped are not uniform—e.g., corn in a given scene may be in various stages of maturity, may have been planted with various row spacings, and also may have been damaged by weather, insects, or poor cultural practices. Thus an object is often multimodally distributed. In addition, such an object must be recognized in the presence of multimodally distributed objects belonging to other classes. Nonetheless, because these multimodal distributions sometimes result from the deterministic errors described earlier, the number of modes can often be reduced by factors of 2 or 3 by correcting such errors.

Changing meteorological conditions can also bring about changes in observed radiances; these result in random variations extending over greater areas and may amount to parameter shifts of 20 to 50%. Such variations make it necessary to correct the estimates of distribution parameters in order to recognize objects successfully.

Two principal forms of adaptation, then, may be applied to multispectral data: First, functions of scanner look angle may be adapted to the scene topography and to changing solar elevation [1]. Second, errors resulting from drifts of many kinds and from changing scene conditions may be adaptively corrected by "tracking" the scene parameters (signature means) as objects are recognized.

This report describes adaptation of the second kind, above—i.e., adaptive modification of scene distribution parameters to improve the recognition accuracy of a classifier.

2.2. AN ADAPTIVE APPROACH

Scene variability as sensed by a multispectral scanner indicates that the radiance of scene objects depends to a considerable extent on natural causes. This is illustrated in Fig. 1 and well described by Kriegler, et al. [2]. These effects can be corrected by pre-processing techniques in which functions of scan angle or functions of spectral bands are used to reduce distribution overlap and thus lessen the sensitivity to these variations. An underlying randomness, however, is not correctable in this manner. Because variations in natural conditions tend to affect collocated objects in the same manner, this randomness is not unstructured—e.g., if the scene irradiance decreases, all objects are proportionately affected (see Fig. 2); or if an instrument bias shift occurs, a similar result is observed. As a result, there is a tendency for a cluster of distributions to preserve its structure when modified by conditions of changing irradiance; this is illustrated in Fig. 3. The parameters of all distributions are affected similarly as a variation occurs and hence can be described as following similar trajectories in spectral space. The object of modeling and analysis is to describe and predict these trajectories. A first approximation is to assume that trajectories follow straight lines which converge to the zero radiance origin of the space (in the noise-free case). For such a trajectory

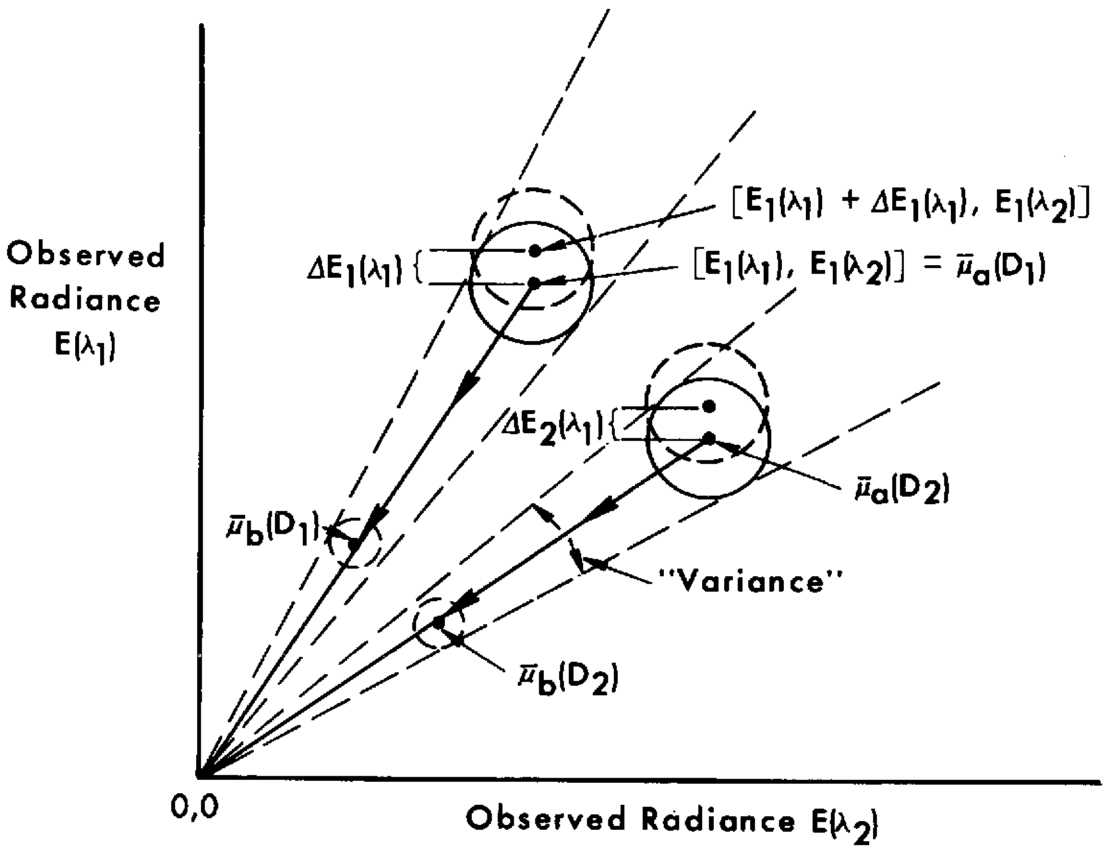


FIGURE 1. POSSIBLE VARIATIONS OF TWO DISTRIBUTIONS (NOISELESS CASE)

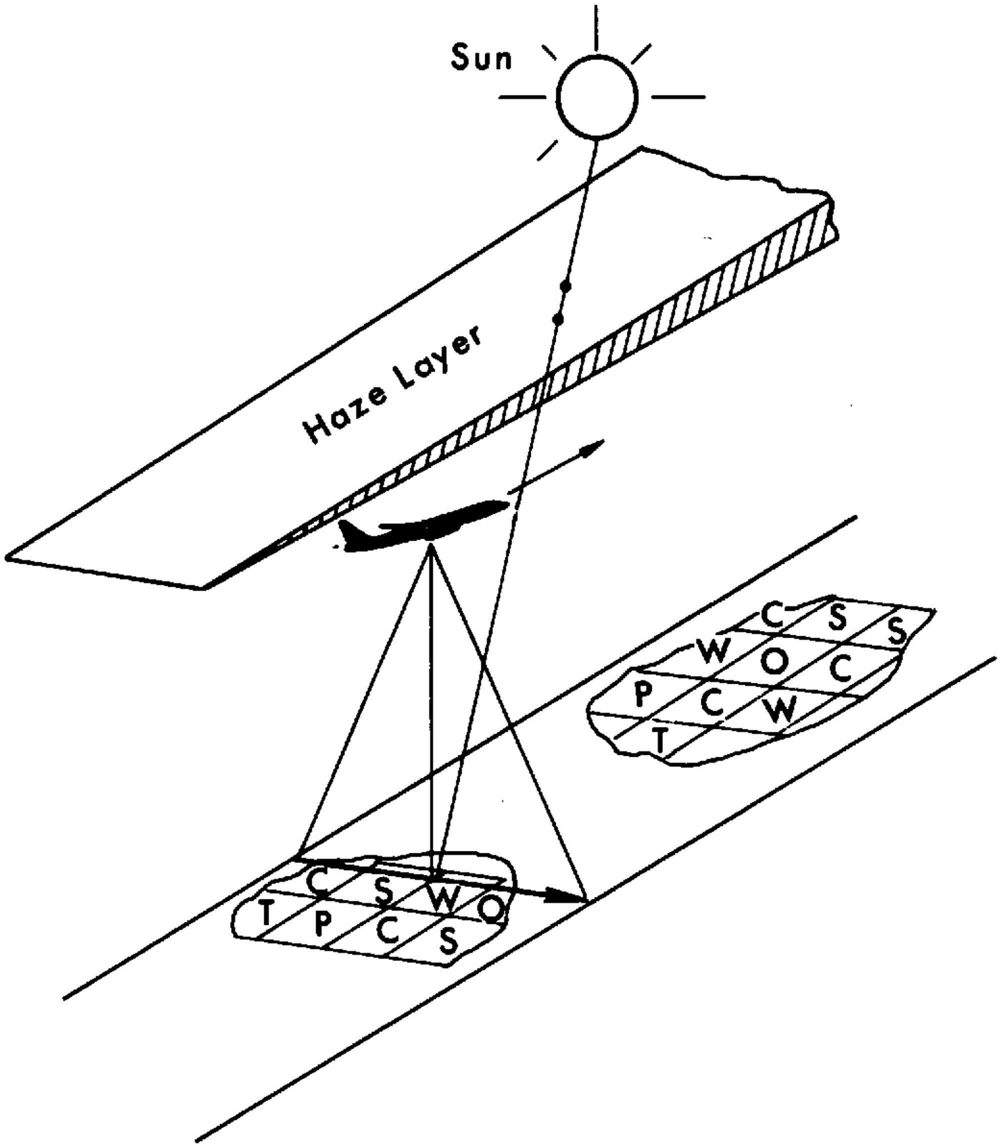


FIGURE 2. A MODEL OF CHANGING IRRADIANCE

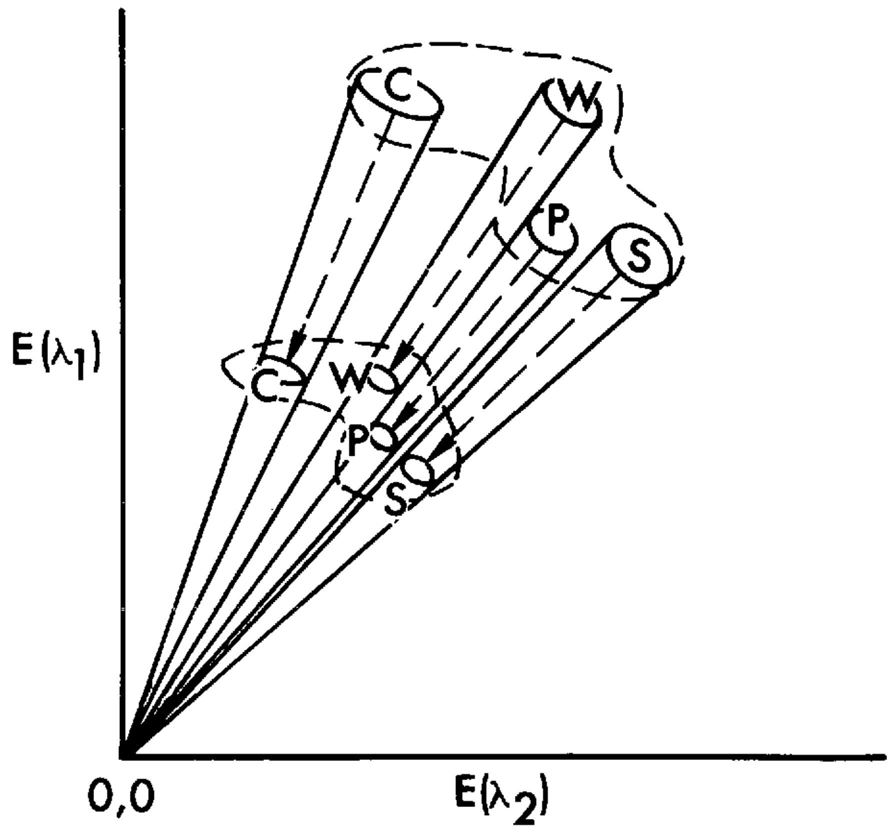


FIGURE 3. EFFECT OF CHANGING IRRADIANCE

the signature distributions would fill a conical space, thus preserving their structure based on the spectral reflectances of the objects observed. Here a spectrally uniform attenuation of irradiance is assumed.

Thus if a classifier can recognize samples of a scene affording some acceptable spatial density for materials of interest, the statistical parameters of all the distributions entered into the classifier may be adaptively modified in two ways: (1) on the basis of each classification and (2) on the basis of predicted trajectories. The spatial density of materials being classified must be such, of course, that the spatial frequency of scene variation is adequately sampled.

Since the problem of theoretically determining the utility of an algorithm based on this model appears formidable, practical experiments were conducted instead to evaluate the various algorithms.

3
ADAPTIVE ALGORITHMS

The task of updating the signatures can be achieved through either of two approaches. One approach is to modify the individual distribution parameters based on changes detected in statistical parameters of the overall scene. Such changes could be detected by suitably averaging, or filtering, the data—as in certain preprocessing techniques now routinely employed. These averaging techniques can best handle systematic variations common to all scene materials; they do not, however, permit the signatures for each scene material to be independently updated. The second approach allows independent updating to be performed without the necessity of additional a priori information (ground truth). Algorithms based on this approach were derived and tested.

Basically, with crop material on the ground scanned by the multispectral sensor at an instant of time, the processing problem is to identify and locate the different ground crops based on the signals generated by the sensing instrument. We assume that the original analog signals have been smoothed and digitized. We further suppose that preprocessing has been performed to eliminate some of the variation caused by scan angle effect. Thus we have signals $y_{u,v}$, $1 \leqq u \leqq U$, $1 \leqq v \leqq V$ representing the output of a multispectral scanner relative to a scene. The u, v coordinates take on integral values and each pair represents a resolution cell of the scene. The u is called the point coordinate and v the line coordinate; y is an m-vector, where m is the number of channels of scanner data being processed. The digitized signals, y, are stored on magnetic tape in sequence according to lowest u coordinate and then lowest v coordinate. It is convenient to put the signals in linear order by

$$X_n = y_{u,v}$$

where

$$n = (v - 1)U + u \text{ and } 1 \leqq n \leqq UV$$

The purpose of processing the data is to classify each data point X_n as being from one of r classes representing different materials in the scene. Since all materials in a scene are not of interest, or because of practical limitations, some class is given the "other" classification and designated the null class. This null class cannot be represented by a convenient distribution. For the classes of interest there is a priori information available about the spectral distribution of signals in each. This information normally takes the form of training sets containing predetermined (known) samples of signals from each class. These training sets are used to estimate the parameters of a Gaussian distribution for each class. Thus each class is associated with a signature distribution.

3.1. NONADAPTIVE PROCEDURE

A typical nonadaptive procedure for classifying signals is as follows. The signal X is classified as material j if

$$f_j(X) \geqq f_i(X) \quad \text{for all } i \tag{1}$$

and

$$Q_j(X) < \tau \tag{2}$$

where f_i is the density of the Gaussian distribution associated with the i-th class and $Q_i(X)$ is the quadratic form occurring in this density function. τ is a threshold parameter. The τ test, required because of the "other" classification, is a chi-square test. Here, if Eq. (2) is not satisfied for any j, the point X is classified as "other."

3.2. ADAPTIVE PROCEDURES

We know that the parameters of the signature distributions vary over the scene; reasons for this variability are given in Section 2.1. Adaptive processing to permit changes in these signature parameters would appear to offer hope for improving classification accuracy. But no extant theory appears directly applicable to our specific problem. So we employed existing theory to guide our choice of adaptive procedures to be tested empirically.

Thus far we have adapted only the means of the signature distributions, utilizing what the literature calls "decision-directed" schemes wherein recognition decisions figure in the update process. To date, we have experimented with four adaptive procedures which, for convenience, we refer to as:

Exponentially weighted running estimates
Exponentially weighted running estimates with interaction
Posterior probability weighted estimates
Posterior probability weighted estimates with interaction

In all of these procedures the data point X is classified by the rules of Eqs. (1) and (2). However, the means A_i, $1 \leqq i \leqq r$, may change from point to point.

3.2.1. EXPONENTIALLY WEIGHTED RUNNING ESTIMATES

With this procedure, if X has been recognized as being in class j, then the present mean A_j of this class is updated to the new value A_j' by the relation

$$A_j' = \left(1 - \frac{1}{W}\right)A_j + \frac{1}{W}X \tag{3}$$

where W is a weighting constant. The term exponential weighting [3] is based on the fact that if $Z_1, Z_2, \ldots, Z_\ell$ are the first ℓ points in the X_n sequence recognized as being in a particular class, and A^k is the k-th updated mean, then A^ℓ is approximated by

$$A^\ell \approx \frac{1}{W}\sum_{k=0}^{\ell} X_k e^{-(\ell-k)/W} \tag{4}$$

where X_0 is the initial value of the mean. As can be seen from Eq. (4) W also corresponds to an exponential decay rate. Abramson and Braverman [4] derive a formula in the form of Eq. (3) for tracking a slowly varying mean of a Gaussian distribution under certain conditions.

3.2.2. EXPONENTIALLY WEIGHTED RUNNING ESTIMATES WITH INTERACTION

Because variation of signature distributions in samples of the same material increases with increased spatial separation, the updating procedure will be less effective if there are long intervals in the X_n sequence which do not represent a particular material. This situation corresponds to having a material located in some part of the scene which is remote from any other concentration of this material in the scene. To deal with this problem, knowledge concerning the trajectory of the set of signature means was used to modify the update procedure. It was assumed that the following ratio remained constant along this trajectory: the ratio of each component of a signature mean to the average of that component over all signature means. Formally,

$$\overline{A} = \rho_i A_i \tag{5}$$

where $\overline{A}$ denotes the average signature mean at any time, A_i is the signature mean of the i-th class, and ρ_i, $1 \leq i \leq r$ is a diagonal matrix obtained by using Eq. (5) after initial signature means A_i have been computed from training sets. In the interactive procedure, not only is the recognized class mean j updated in Eq. (3) but also each A_i, $i \neq j$, is then updated by the formula

$$\overline{A}' = \rho_i A_i' \quad i \neq j \tag{6}$$

where

$$\overline{A}' = \rho_j A_j' \tag{7}$$

3.2.3. POSTERIOR PROBABILITY WEIGHTED ESTIMATES

In discussing this procedure it is convenient to first define the following:

A_i, $1 \leqq i \leqq r$ = the present value of the mean of i-th signature

M_i = covariance matrix of i-th signature

X = present observed point

$$L_i(X) = \frac{1}{|M_i|^{1/2}} e^{-1/2 \left\langle X-A_i, M_i^{-1}(X-A_i) \right\rangle}$$

$$R_i(X) = \frac{L_i(X)}{\sum_j L_j(X)}$$

$$Q_i(X) = \left\langle X - A_i, M_i^{-1}(X - A_i) \right\rangle$$

$\langle u, v \rangle$ = inner product of the vectors u and v

Now if the a priori probability that X belongs to i-th class is $1/r$, then $R_i(X)$ represents the posterior probability that X belongs to class i. Thus $R_j(X)$ is a measure of the confidence we have in correct identification after X has been classified by Eqs. (1) and (2). This measure of confidence may be reflected in determining the weight used in updating the means. Here is one way of applying such a confidence measure: The point X is classified in class j if

$$L_j(X) \geqq L_i(X) \quad 1 \leqq i \leqq r \tag{8}$$

and

$$Q_j(X) < \tau_1 \tag{9}$$

If Eq. (9) is not satisfied for any j, the point X is classified as "other." A_j is updated if in addition to Eqs. (9) and (10)

$$Q_j(X) < \tau_2 \tag{10}$$

Then the updated value A_j' of A_j is

$$A_j' = 1 - \frac{R_j(X)}{W} A_j + \frac{R_j(X)}{W} X$$

We call τ_1 the recognition parameter and τ_2 the update parameter. Then because we usually require a more stringent condition to be satisfied for update than for recognition, we specify

$\tau_2 < \tau_1$. The decay rate parameter W is retained to provide more flexible control over the influence of observation X on the present mean. It is seen that the a priori probabilities of the occurrence of any class are taken to be equal. Subsequent investigation, however, may lead to modification of this premise.

3.2.4. POSTERIOR PROBABILITY WEIGHTED ESTIMATES WITH INTERACTION

For reasons given in Section 3.2.2, the interactive procedure was appended to the preceding procedure by using Eqs. (6) and (7) when the conditions of Eqs. (8), (9), and (10) are satisfied for some j.

3.3. REMARKS ON ADAPTIVE PROCEDURES

The four updating procedures just described provide a basis for continuing investigation of adaptive processing as it may be applied to multispectral data. Some possible modifications are apparent: use of a priori probabilities of class occurrence; use of conditional probabilities to take into account that the point X_n is likely to be in class j if the preceding point X_{n-1} was in class j (especially in agricultural scenes); and updating of covariance matrices of signatures as well as the means. It must be borne in mind that inclusion of these modifications would increase processing time.

4
EXPERIMENTAL RESULTS

We now present the results of experiments using the algorithms defined in Section 3. For purposes of comparison these experiments were performed with scanner data from the same scene. These data were gathered on a flight near Lafayette, Indiana at 1400 hours on June 30, 1966, at an altitude of 3500 ft. The associated ground-truth map appears in Fig. 4.

4.1. NONADAPTIVE PROCESSING

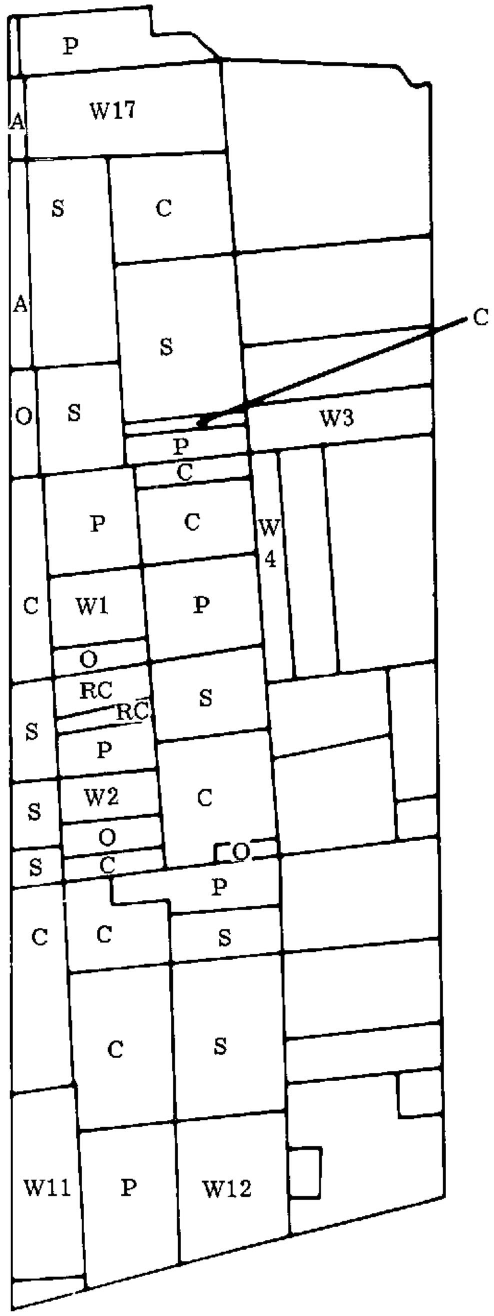
A study and experiment was performed previously [5] on recognizing wheat in this scene which was at first thought to be well described by non-time varying statistics. This wheat recognition was satisfactory over the length of the run (about 4 miles) because this result was actually obtained with two distributions describing wheat, one at each end of the run and designated in Fig. 5 as Field A and Field B. In addition the recognition was affected by radiance variations along the scan line. Figure 5a shows the results of recognition processing using a single distribution, that of Field A, to recognize wheat; a gradual deterioration in recognition will be noted over the run from South (bottom) to North (top). Reasons for this were not clear at the time, but it was postulated that an atmospheric effect similar to that shown in Fig. 2 may have existed.

An investigation of the data from these two fields indicated that a large shift occurred in the means (Δ mean) of the distribution. This is summarized in the following table:

Wavelength (μm)	.44-.46	.48-.50	.55-.58	.62-.66	.72-.80
Δ Mean (volts)	.25	.44	.35	.499	.545
Field A - Std. Deviation (volts)	.086	.146	.100	.125	.160
Δ Mean/σ	2.9	3.06	3.5	4.0	3.4

4.2. ADAPTIVE EXPERIMENTS—ONLY WHEAT SIGNATURE MEAN UPDATED

The initial adaptive experiment [6, 7] was based on the procedure described in Section 3.2.1 with the modification that only the wheat signature was updated. W was set at 300 data points. We actually used the knowledge that there were about 4800 wheat points in the scene and that about 3000 were recognized with the nonadaptive scheme. The W figure of 300 appeared small enough to permit suitably rapid modification of the wheat signature. This initial adaptive experiment produced the results shown in Fig. 4b. The improvement in recognition is apparent, actually showing an increase from 186 acres to 252 acres, which gives an overall accuracy of 83% vs. 62%. Improvement is even more dramatic if the training field is not included in the computations.



LEGEND

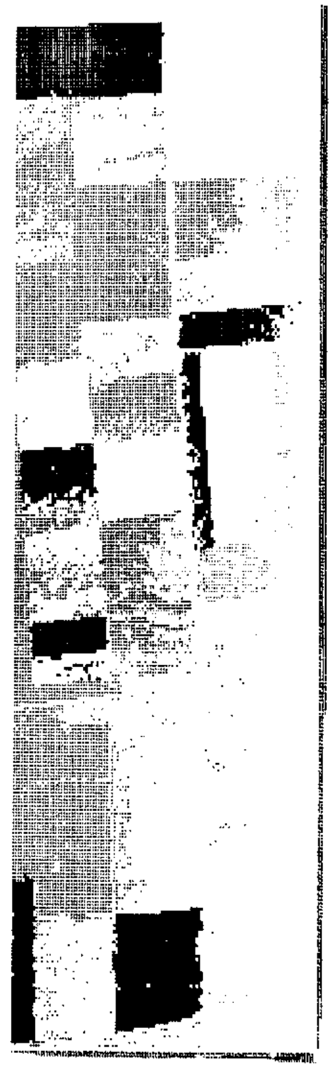
W Wheat
S Soybeans
C Corn
A Alfalfa
P Pasture
RC Red Clover
U Unknown

(Fields Not Labeled are Unknown Crops)

Flight Direction

FIGURE 4. GROUND TRUTH MAP FOR DATA COLLECTED NEAR LAFAYETTE, INDIANA ON 30 JUNE 1966

(a) Fixed Signature



(b) Updated Signature

FIGURE 5. COMPARISON OF WHEAT RECOGNITION RESULTS FOR FIXED AND ADAPTED SIGNATURES

The results can be further analyzed by comparing the recognition map in Fig. 5b with the ground "truth map" of Fig. 4. Some false classification occurs in two oat fields just below the two small fields located in the left center of the map; this amounted to 30 samples in the lower field and approximately 100 samples in the upper field. Since only the means of wheat were being adapted rather than those of other distributions (corn, soybeans) some capture of the process was expected. The fixed signature produced 13 samples in the lower field and about 35 samples in the upper field. Some capture occurred.

Since the spatial distribution of wheat is not uniform, one would expect the angular effect to be also reduced; this reduction can be observed, particularly in the upper right-hand field which was known to have a large angular dependence.

A simple experiment was also conducted to determine the ability of the procedure to overcome a transient error. With the wheat mean vector obtained at the end of the run used as an initial estimate,a classification run was made. Results are shown in Fig. 6. No discernible difference exists after the first 600 wheat points, which is twice the weighting constant. A difference of 202 elements exists between the two adaptive runs. The estimates of the means agreed to the fourth decimal place at the end of the run.

4.3. ADAPTIVE EXPERIMENTS —WHEAT, CORN, AND SOYBEANS UPDATED

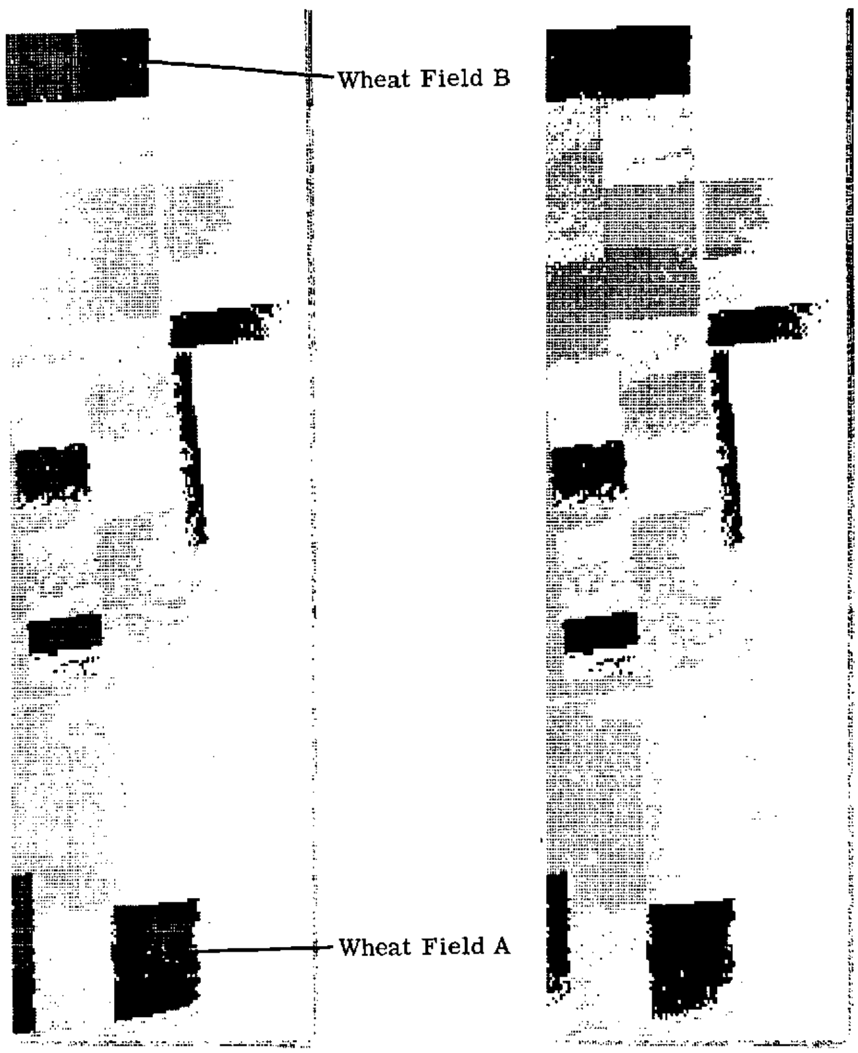
Results of our experiment in adapting only the wheat signature led to the belief that results would be improved by also adapting other signatures. In the next set of experiments the wheat, corn, and soybean means were updated using the noninteractive procedure of Section 3.2.1. The threshold parameter, τ, of Eq. (2) was set at 25, but the decay rate, W, was varied from 100 to 1000. Results showed a tendency for the soybean signature to capture pasture.

We also tried the interactive algorithm described in Section 3.2.2. Results showed a tendency for the corn signature to capture soybeans. This capture effect was pronounced on one side of the scene along the flight line and evident in both the interactive and noninteractive experiments.

4.4. EXPERIMENTS WITH REDIGITIZED DATA

At this stage it appeared advisable to redigitize* all 12 spectral channels of the analog data and correct for the dynamic dark level shifts. Angle correction procedures were applied

*Our reason for redigitizing was that the original digitization had been done on only 6 nonadjacent spectral bands for only the video portion of the data; the dark level and calibration portions were not digitized. Thus, no channel selection procedure was at that time employed nor could dark level corrections be made. All subsequent adaptive experiments were performed on the digitized 12-channel data.



(a) Initial Signature From Wheat Field A

(b) Initial Signature Using Previous Update

FIGURE 6. WHEAT RECOGNITION RESULTS SHOWING THE CONVERGENCE OF THE ADAPTED SIGNATURE

first. Then spectral channel selection techniques were utilized to select a subset of the six spectral channels exhibiting best separation between the corn and soybean signatures (which tend to be close to each other).

In the next set of experiments, the parameter, τ, and the decay rate, W, were varied. For these experiments the pasture signature was updated as well as those of soybeans, corn, and wheat. Based on our previous experience with capture of one signature class by another class, tighter (smaller variance) signatures were desired and obtained. One reason for this reduced variance had to do with application of the angle correction techniques. Another reason was that anomalies in the training fields were avoided when initial signatures were calculated.

Figure 7 shows results for the exponentially weighted noninteractive and interactive processes. W was set at 300 and τ was set at 25 for both these experiments.

Some problems were immediately evident. In the noninteractive case a capture process seemed to be taking place, thus causing distributions to be confused with each other as processing progressed. Also, the percentage of wheat detected was low.

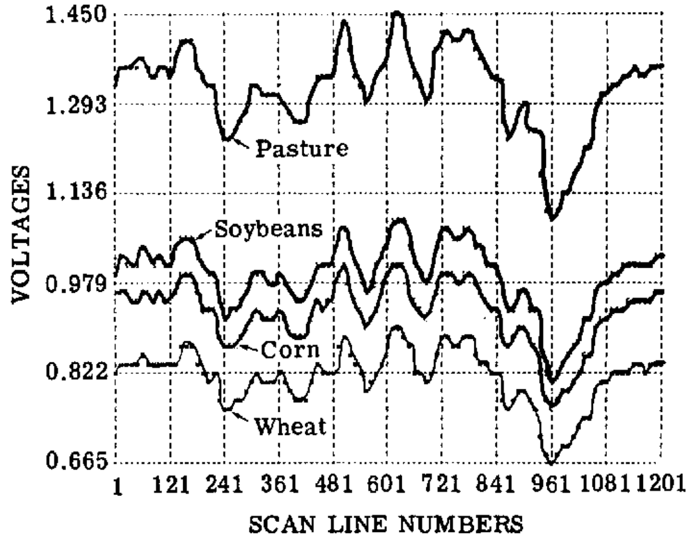
In the interactive process signatures are captured in a different manner. As data processing continues, variations resulting from more frequently occurring distributions induce a motion of the mean structure which prevents recognition and adaptation by distributions occurring less frequently. In other words, distributions (such as wheat) occur so infrequently that the forced (interactive) mean structure variation does not adequately reflect the actual change of such a distribution.

The means as a function of distance along the flight path are plotted in Fig. 8 for one of the near-infrared spectral bands. Here, for the case of the noninteractive algorithm (Fig. 8b) several crossovers occur in the mean plots. This type of crossover occurred for each of the several wavelengths employed. The mean plot for the interactive technique (Fig. 8a) indicates that the distribution for objects with greatest spatial extent tends to control the means; this results in poorer recognition of objects having lesser spatial extent.

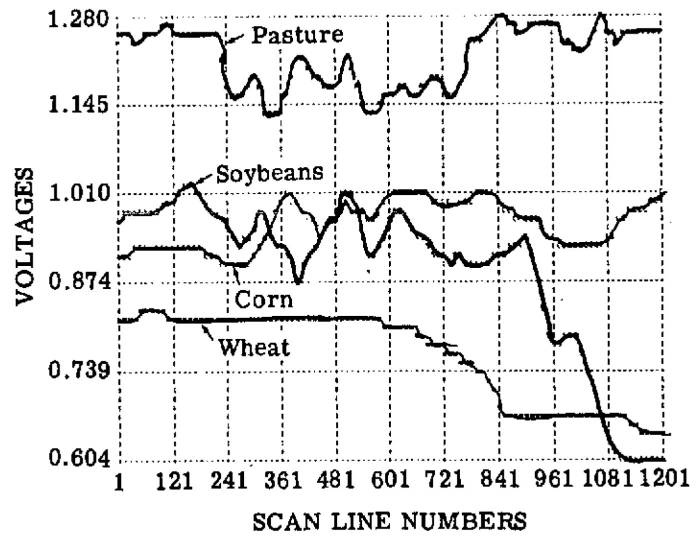
To improve the recognition percentage and minimize captures, the rejection limit (threshold parameter) was increased to $\tau = 100$ and the weighting term to $W = 1000$. Figures 9a and 9b show results for the noninteractive and interactive cases, respectively. These results are satisfactory except that toward the end of the run the classifications of corn and soybeans tend to be confused. This problem may be analyzed by inspecting the mean plots shown in Fig 10. In the noninteractive plots (Fig. 10b) which are substantially correct toward the end of the run, we see that the soybean mean shows considerable variation, actually crossing those of corn and wheat in some spectral bands. Since this cannot happen using the interactive algorithm, recognition errors may have occurred. Note that the wheat field in the lower left corner is missed when the parameter $\tau = 25$

(a) Interactive (b) Noninteractive

FIGURE 7. COMPARISON OF RECOGNITION RESULTS FOR INTERACTIVE AND NONINTERACTIVE UPDATE OF SIGNATURES. (Exponentially weighted running estimates with W = 300, τ = 25.)

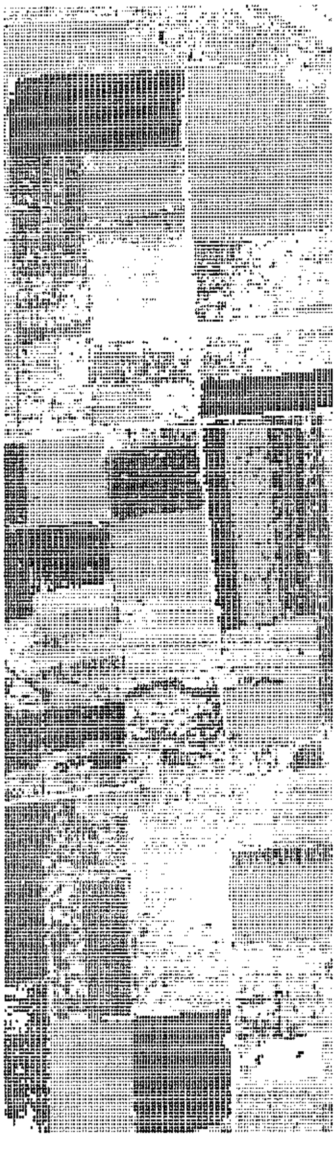


(a) Interactive



(b) Noninteractive

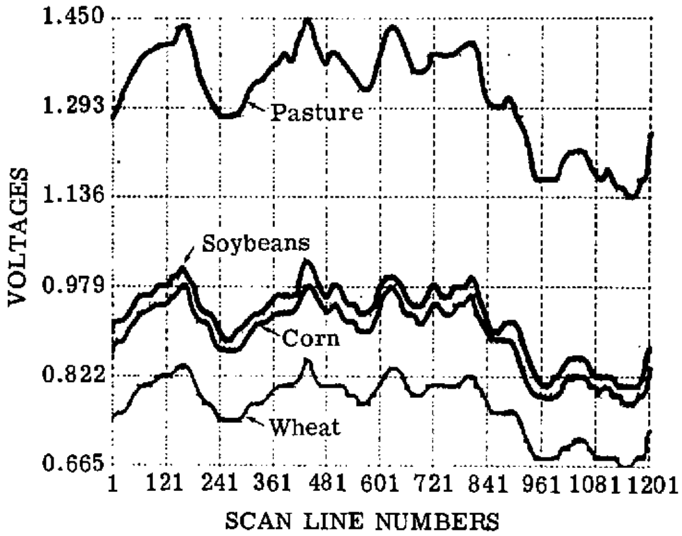
FIGURE 8. PLOT OF MEAN VALUE VARIATIONS ALONG THE FLIGHT PATH. (Exponentially weighted running estimates with W = 300, τ = 25.)



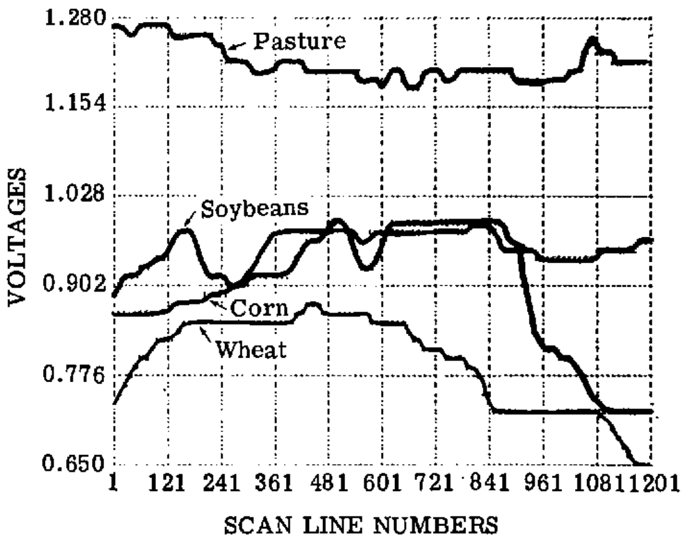
(a) Interactive

(b) Noninteractive

FIGURE 9. RECOGNITION RESULTS FOR IMPROVED UPDATE PARAMETERS. (Exponentially weighted running estimates with W = 1000, τ = 100.)



(a) Interactive



(b) Noninteractive

FIGURE 10. MEAN VALUE VARIATIONS ALONG THE FLIGHT PATH FOR IMPROVED UPDATE PARAMETERS. (Exponentially weighted running estimates with W = 1000, τ = 100.)

(see Fig. 7), but the same field is picked up for $\tau = 100$ (Fig. 9). Also note the poor recognition of wheat field B in Fig. 7, and the substantially improved recognition of this field in Fig. 9.

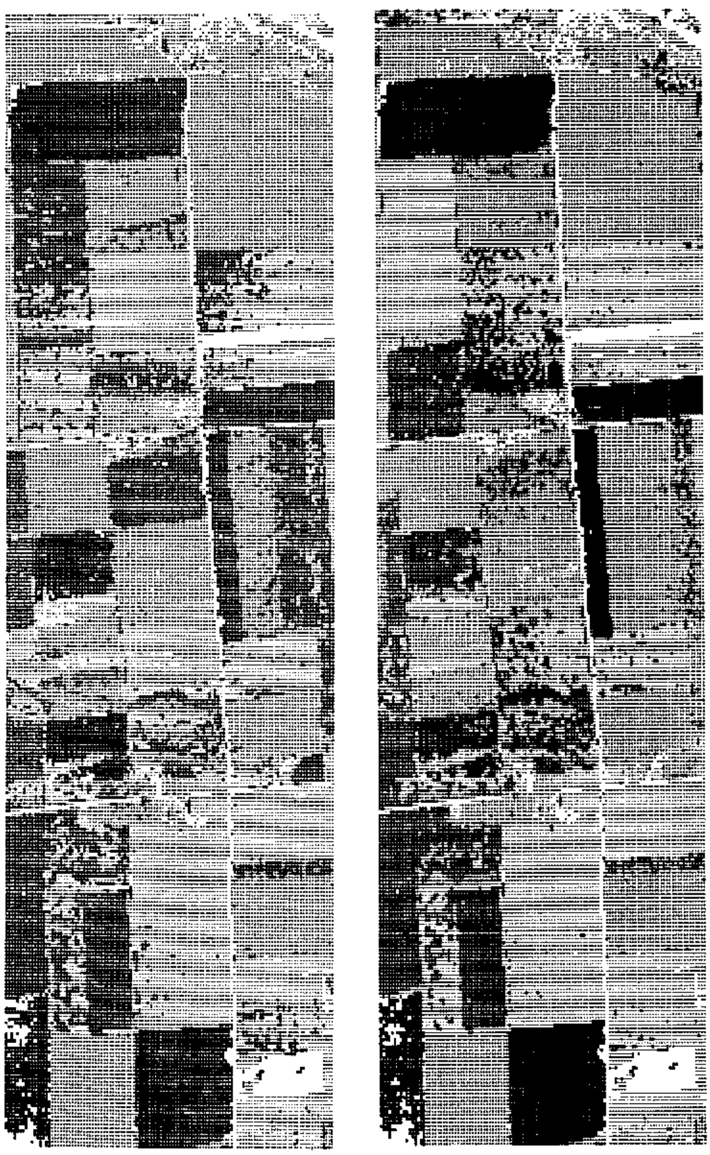
For experiments using parameter values $\tau = 25$ and W = 150, results (not illustrated) were qualitatively similar to the results obtained in using the values $\tau = 25$ and W = 300 (Fig. 7) for both the noninteractive and interactive cases.

Results with parameter settings $\tau = 60$ and W = 1000 (not illustrated) were superior to results for $\tau = 25$, W = 300 (Fig. 7) but inferior to results for $\tau = 100$, W = 1000 (Fig. 9). Only about 8 to 10% of the wheat field in the lower left hand corner of the scene was recognized for the noninteractive and interactive cases.

Runs were then made using the adaptive procedures of Section 3.2.3 and Section 3.2.4. (Both of these procedures are based on posterior probability weighted estimates, the second being the interactive version of the first.) Recognition maps for parameter settings $\tau_1 = \tau_2 = 100$ and W = 1000 are shown in Figs. 11a and 11b for the noninteractive and interactive cases, respectively. Comparison of these maps* with Figs. 9a and 9b (the corresponding runs for exponentially weighted running estimates) shows only negligible differences. Some difference can be seen within the boundaries of rectangular fields where there is a large degree of confusion between corn and soybeans. For corn recognition the posterior probability weighted estimates offer a slight improvement over the exponentially weighted estimates. Mean plots for these runs are also almost identical. From this evidence one can infer that the value of $R_j(X)$ is close to 1 when the mean A_j is updated. It has been observed on the SPARC processor that usually one $L_i(X)$, say $L_j(X)$, is much larger than all the others; this would result in $R_j(X)$ being close to 1.

When the decay rate W was increased to 2000, results for the noninteractive case of posterior probability weighted estimates became degraded over those for W = 1000, but in the interactive case there was little change. For the interactive case degradation approaching the nonadaptive results occurred for W = 4000. The reason for this difference is that in the interactive case the signatures are updated more frequently.

*To identify corn and soybean fields, see the ground-truth map (Fig. 4).



(a) Interactive (b) Noninteractive

FIGURE 11. RECOGNITION RESULTS FOR POSTERIOR PROBABILITY WEIGHTED ESTIMATES. (W = 1000, $\tau_1 = \tau_2 = 100$.)

5
CONCLUSIONS AND RECOMMENDATIONS

It has been demonstrated that simple adaptation of the signature means yields greatly improved recognition performance over the usual (nonadaptive) schemes for processing multispectral data. Previous attempts to compensate for varying signatures over a scene resulted in the characterization of a single material by either a multimodal distribution or several unimodal distributions. That approach requires a number of training sets for each material distributed over the scene to permit signature estimation. Also, great effort and expense is required to gather ground truth information, so this approach does not appear to be rewarding. In addition, processing time is increased in proportion to the number of modes of signature distribution, or to the number of distributions used to represent a single material.

Generally, the noninteractive procedure yielded somewhat superior results over corresponding interactive procedures, although in some instances the interactive procedure assisted in resolving confusion between corn and soy. The present interactive approach assumes unity correlations between the trajectories of the various signatures. This does not permit crossover of components of different signature means. However, the crossover phenomenon still occurs. It would be desirable to experiment with interactive schemes of less than unity correlation between trajectories that would permit crossovers. Further, interactive schemes should be devised where interaction between the trajectories of the means is not limited to interaction between corresponding spectral components.

No significant difference was found between the effectiveness of exponentially weighted running estimates and that of posterior probability weighted estimates. This is because the posterior probability $R_j(s)$ is usually close to 1. Consequently, the points X which cause the greatest changes in the mean upon update are those which are farthest away, and thus have the highest probability of being classified incorrectly. Other procedures which use weights to express confidence in the correctness of classification should be tested. The posterior probability weight scheme is that sort of scheme, but in practice the distribution of $R_j(x)$ turns out to be too flat.

Although preprocessing techniques reduce angular effects in the data, these effects remain troublesome. One can view the variation in signatures as being two-dimensional, one dimension along the flight line and the other along the scan line. It would be desirable to test adaptive schemes that explicitly handle both of these effects. In order to achieve these results more effectively, the adaptive techniques should be based on a more theoretical framework rather than to continue the more empirical method of conducting adaptive experiments. It may be possible to mold the Kalman-Bucy theory into a decision-directed scheme for this purpose [8]. However, application of this theory usually entails enormous computation and this aspect might neutralize its usefulness in multispectral processing.

Up to the present, the adaptive experiments have all been performed on a single data set. It is recommended that the adaptive schemes be tested on many other data sets collected over extended areas to determine whether results obtained to date are universal in nature. The processing of many such data sets on a conventional digital computer can take a prohibitively long time and be very expensive. The best mechanization for performing these experiments may be a hybrid processor [9].

REFERENCES

1. Kriegler, F. J., "Implicit Determination of Multispectral Scanner Data Variation Over Extended Areas," Report No. 10259-1-X, Proceedings of the Seventh International Symposium on Remote Sensing of Environment, May 1971, at The University of Michigan, Ann Arbor.

2. Kriegler, F. J., et al., "Preprocessing Transformations and Their Effects on Multispectral Recognition," Proceedings of the Sixth International Symposium on Remote Sensing of Environment, October 1969, at The University of Michigan, Ann Arbor.

3. Trimble, C. R., "Calibrated Real-Time Signal Averaging," Hewlett-Packard Journal, April 1968, p. 11.

4. Abramson, N. and D. Braverman, "Learning to Recognize Patterns in a Random Environment," IRE International Symposium on Information Theory, Vol. IT-8, pp. 58-63, 1962.

5. Marshall, R. E., et al., "Use of Multispectral Recognition Techniques for Conducting Rapid, Wide-Area Wheat Surveys," Proceedings of the Sixth International Symposium on Remote Sensing of Environment, October 1969, at The University of Michigan, Ann Arbor.

6. Marshall, R. E., F. J. Kriegler and W. Richardson, "Adaptive Multispectral Recognition of Wheat," Willow Run Laboratories Report No. 31650-85-S, presented at the 1971 IEEE Decision & Control Conference, 15-17 December 1971, at Miami.

7. Marshall, R. E., F. J. Kriegler and W. Richardson, "Adaptive Multispectral Recognition of Wheat Fields," Willow Run Laboratories Report No. 31650-86-S, presented at the EAI Symposium on Automatic Photointerpretation, 7-8 December 1971, at Washington, D.C.

8. Bucy, R. S., "Linear and Nonlinear Filtering," Proceedings of the IEEE, Vol. 58, No. 6, June 1970, pp. 854-864.

9. Marshall, R. E. and F. J. Kriegler, "Multispectral Recognition Techniques: Present Status and a Planned Hybrid Recognition System," presented at the Ninth IEEE Symposium on Adaptive Processes, Decision and Control, December 1970, at University of Texas, Austin; published in Proceedings, Vol. XI, pp. 2.1-2.9.

SELECTED BIBLIOGRAPHY

Bishop, A. B., "Adaptive Pattern Recognition," Wescon Technical Papers, 1963.

Booker, A. H., C. Y. Ong, J. P. Burg, and G. D. Hair, "Multiple Constraint Adaptive Filtering," Texas Instruments, SCI Services Division, Dallas, April 1969.

Braverman, D., "Learning Filters for Optimum Pattern Recognition," IRE Trans., P017-8, July 1962.

Bryan, J. S., "Experiments in Adaptive Pattern Recognition," IRE Trans. Mil. Electron, Vol. MTL-7, April-July 1963, pp. 176-179.

Chow, C. M. and C. W. Liu, "An Approach to Structure Adaption in Pattern Recognition," IEEE Trans. Sys. Sci. Cybern., Vol. SEC-2, No. 2, December 1966.

Cooper, D. B., "Adaptive Pattern Recognition and Signal Detection Using Stochastic Approximation," IEEE Trans. Electron. Comput., Vol. EC-13, No. 3, June 1964, pp. 306-307.

Fukunaga, K., Introduction to Statistical Pattern Recognition, New York: Academic Press, 1972.

Glanz, Filson W., "Statistical Extrapolation in Certain Adaptive Pattern Recognition Systems," Stanford University, Stanford, May 1965, AD 467 627.

Griffiths, L. J., "Signal Extraction Using Real Time Adaption of a Linear Multichannel Filter," Technical Report No. TR 67881-1, Stanford Electron. Labs, Stanford, DOC SEL-60-017, February 1968.

Groner, G. F., "Statistical Analysis of Adaptive Linear Classifiers," Technical Report No. 6761-1, Stanford Electron. Labs, Stanford, Report No. SEL-66-026, April 1964.

Hilborn, Charles G., and Demetrios G. Lainiotis, "Optimal Unsupervised Learning Multicategory Dependent Hypotheses Pattern Recognition," IEEE Trans. Inform. Theory, Vol. IT-14, No. 3, May 1968.

Ho, Y. C. and A. K. Agarwala, "On Pattern Classifications Algorithms—Introduction and Survey," Proc. IEEE, Vol. 56, December 1968, pp. 2101-2114.

Kaimieraczak, W. and K. Steinbuch, "Adaptive Systems in Pattern Recognition," IEEE Trans. Electron. Comput., 1963.

Kanefsky, M. and J. B. Thomas, "On Adaptive Nonparametric Detection Systems Using Dependent Samples," IEEE Trans. Inform. Theory, Vol. IT-11, October 1965, pp. 521-526.

Koford, J. S. and G. F. Groner, "The Use of an Adaptive Threshold Element to Design a Linear Optimal Pattern Classifier," IEEE Trans. Infor. Theory, Vol. IT-12, January 1965, pp. 42-50.

Lainiotis, D. G., "Sequential Structure and Parameter Adaptive Pattern Recognition, Part I: Supervised Learning," IEEE Trans. Infor. Theory, Vol. IT-16, September 1970.

Martinez, Hugo N., "An Introduction to the Theory of Adaptive Pattern Recognition," AD 600 157.

Mattson, R. L., "The Design and Analysis of an Adaptive System for Statistical Classification," M.S. Thesis Elec. Engrg. Dept., Massachusetts Institute of Technology, May 1959.

Moschner, J. L., "Adaptive Filtering with Clipped Input Data," Technical Report No. TR 6796-1, Stanford Electron. Labs, Stanford, DOC SEL-70-053, June 1970.

Mucciardi, Anthony N., "Adaptive Pattern Recognition Using Nonlinear Elements," AD 623 215.

Patrick, F. H. and J. C. Hancock, "Nonsupervised Sequential Classification and Recognition of Patterns," IEEE Trans. Infor. Theory, Vol. IT-12, No. 3, 1966.

Patterson, J., I. Wagner, and B. Womack, "A Performance Criterion for Adaptive Pattern Classification Systems," 1966 Proc. Joint Automatic Control Conf., pp. 38-46.

Patterson, J. D., An Adaptive Pattern Classification System, Ph.D. Dissertation, University of Texas, 1966.

Patterson, J. D. and B. F. Honack, "An Adaptive Pattern Classification System," IEEE Trans. Sys. Sci. Cybern., Vol. SEC-2, No. 1, August 1966.

Rajaskarah, P. K. and M. D. Srinath, "Structure and Parameter Adaptive Pattern Recognition with Supervised Learning: A New Formulation," IEEE Trans. Infor. Theory, July 1971.

Roberts, L. G., "Pattern Recognition with an Adaptive Network" IRE International Convention Record, Pt. 2, 1960, pp. 66-70.

Sammon, T. W., "Interactive Pattern Analysis and Classification," IEEE Trans. Comput., Vol. C-19, July 1970, pp. 594-615.

Saradis, G. N., Z. J. Nikolic, and K. S. Fu, "Stochastic Approximation Algorithms for System Identification, Estimation and Decompositions of Mixtures," IEEE Trans. Sys. Sci. Cybern., Vol. SSC-5, January 1969, pp. 8-15.

Sears, R. B., Jr., "Adaptive Representations for Pattern Recognition," IEEE Trans. Sys. Sci. Cybern., Vol. SEC-1, No. 1, November 1965.

Sebestyen, G. S., "Pattern Recognition by an Adaptive Process of Sample Set Construction," Proc. IRE Professional Group on Information Theory, IT-B, 5, September 1962.

Van Trees, H. L., Detection, Estimation and Modulation Theory: I., New York: Wiley, 1968.

Wadner, Terry J., "A Performance Criterion for Adaptive Pattern Classification Systems," Dept. of Electrical Engineering, University of Texas, Austin, 1964.

Walter, C. N., "Computer Aided Methods for Monitoring Performance of Class of Adaptive Pattern Attribute-Organizing Procedures, IEEE Nenen Record, 1963.

Walter, C. M., "Some Comments on Interactive Systems Applied to the Reduction and Interpretation of Sensor Data," In IEEE Computers and Communications Conf. Rec. (IEEE Spec. Pub. 69 C67-MVSEC), 1969, pp. 109-112.

Wee, W. G., "Generalized Inverse Approach to Adaptive Multiclass Pattern Classification," IEEE Trans. Comput., Vol. C-17, 1968, pp. 1157-1164.

Widrow, B., "Adaptive Sampled Data Systems—Statistical Theory of Adaption," 1959 Wescon Convention Record, Part 4.